Wind Speed Forecast Based on Support Vector Machine

Yang Xiao-hong[1,a] ,Tang Fa-qing[1,b]

[1]School of Energy and Power and Mechanical Engineering, North China Electric Power University, Baoding 071003, China

[a] yangxh@ncepu.edu.cn , [b]1322960631@qq.com

Key words: support vector regression; wind speed; forecast

Abstract. Environmental pollution, climate change, energy tension is now the global prominent question, the development and utilization of renewable energy to become one of the key measures to solve these problems. Wind power acts as a kind of safety, environmental protection, clean, abundant renewable energy by all countries focusing on and extensive using. The uncertainty of wind power makes high demands for operation scheduling, and more accurate wind speed forecast is of great significance to the effective use of wind power. Predictions based on support vector regression machine of wind speed were new application in recent years, it has received the widespread attentions. Using support vector regression machine predict the wind speed of a wind farm data, the analysis results show that the method of support vector regression machine can more accurately predict the wind speed and also proves the superiority of the prediction of wind speed.

Introduction

Wind power generation is the main form of utilization of wind energy. Recent years, with wind power technology continues to mature and large-scale wind farms have been expanding, greatly improving the utilization of wind energy. However, the wind speed with a random fluctuation, indirect characteristics and susceptible to season, climate and other factors, and therefore to develop wind farms scheduling, efficient use of wind energy has brought great difficulties. Coupled with the rapid development of China's wind power industry and continuous improvement of the installed capacity, therefore, a more accurate prediction method of wind speed is particularly important for the development planning, rational use of wind energy and reduce operation cost of the system [1].

There are a lot of methods to predict wind speed now, such as artificial neural network, fuzzy logic method, numerical weather prediction method etc [2], but the method of support vector regression machine to forecast wind speed takes its unique advantages to let people shines [3].

Support Vector Regression (SVR) is based on statistical theory. It is a perfect combination of statistical theory of VC dimension theory, promotion of community and structural risk minimization theory and kernel, the optimal regression hyperplane theory and SVM regression theory. Because of it takes the minimum confidence range as the optimization objective, and it takes the training error as the condition of the optimization problem, the generalization ability of the support vector regression machine is better than lots of traditional methods of wind speed forecast. Another great advantage of support vector machine is that it can turn the problem of solving the optimal return hyper plane into the problem of solving the quadratic program which is largely to ensure the uniqueness of solution, easy to get the global optimal solution and also to ensure the accuracy of the forecast [4].

Support Vector Regression Machine Theory

Referring to support vector machine regression analysis, we aim to find a plane which is called the superior regression flat surface, to make the best fit to the training data. The superior regression flat surface requires that the total deviation of all sample points from the hyper plane is minimum, and the sample points are all within the two boundary lines [5].
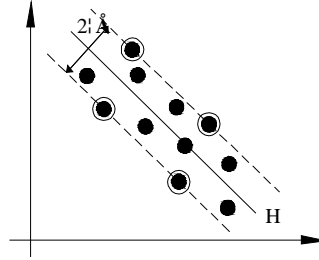

Fig.1   optimal regression hyperplane

According to the Cover theorem, the sample data of the input space is mapped to the feature space with high dimension and infinite dimension by nonlinear function transformation, and the linearly separable probability of these samples is higher in feature space. At the same time, it brings a problem that when the nonlinear function is mapped to the high dimension space, the complexity of the algorithm will be increased significantly, and it is easy to cause the "curse of dimensionality" of the algorithm. The existence of the kernel function is a good solution to this problem. The computation of kernel function does not depend on the dimension of feature space, and it is only concerned about the dimension of input space, so the computational complexity is not increased after nonlinear transformation, which is why the kernel function can overcome the problem of "Curse of dimensionality" [6].
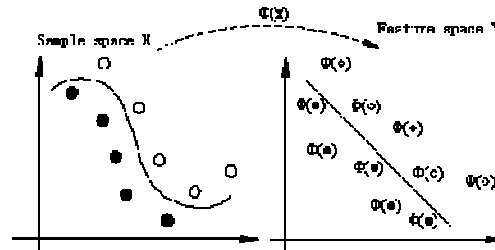

Fig.2   nonlinear transformation

Different kernel functions constitute different support vector machine. As long as the function satisfies the Mercer theorem, it can be used as a kernel function [7].

Mercer theorem: a necessary and sufficient condition for the kernel function $K(x, y)$ to satisfy $f(x) \cdot f(x_i) = K(x, x_i)$: Any function g (x) satisfying formula (1), the formula (2) is established.

$$\int g(x)^2 dx < \infty. \tag{1}$$

$$\iint k(x, y)g(x)g(y)dxdy \geq 0. \tag{2}$$

In support vector machine, the commonly used kernel functions [8] are the following:

(1)Linear kernel function

$$K(x, y) = x \cdot y. \tag{3}$$

(2)P order polynomial kernel function

$$K(x_i, x_j) = (x_i \cdot x_j + 1)^p, \; p = 1, 2, \mathbf{L} . \tag{4}$$

(3)Gaussian Radial Basis Function kernel

$$K(x, x') = \exp(\frac{\| x, x' \|}{s^2}). \tag{5}$$

(4)Sigmoid inner product function

$$K(x_i, x_j) = \tanh(c_1(x_i \cdot x_j) + c_2). \tag{6}$$

The problem to be solved by regression analysis is: According to the limited training samples seek the implicated regression function, and then use the obtained regression function to predict the future data. The support vector machine method based on kernel function is proposed to find a new way to solve the nonlinear regression problem. In this way, we should define the optimal return hyper plane at first, and then solve the linear SVR problem; finally, we should introduce the kernel function to solve the nonlinear SVR problem.

Experimental research and analysis

In order to study the application of support vector regression machine in wind speed prediction, this paper selects a wind field data to predict the wind speed. The wind speed of each sample point is selected as the target value. The pressure, the angle and the temperature of each time point are used as the eigenvalues, and the sampling interval is 10 minutes.

We select 300 sample points of wind speed data as training data, and we predict the wind speed data from 301 to 310.
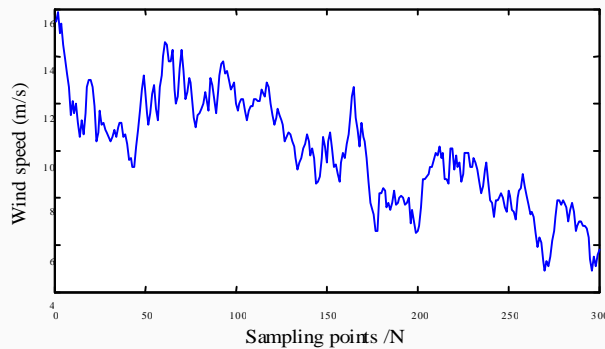

Fig.3 waveform figure of training data to wind speed

In the training model, we choose the more commonly used radial basis kernel function as the kernel function. The radial basis kernel function is a universally applicable kernel function, which can be applied to any distribution of samples by the choice of parameters. The biggest feature of the radial basis function model is: by Euclidean distance, it transforms a multidimensional problem into a one dimensional problem with Euclidean distance as the independent variable. It is currently the most widely used in support vector machines. Because the radial basis kernel function has a strong learning ability, and the parameter selection is simple. Therefore, in this paper, the model is selected as the radial basis kernel function.

Whether the generalization performance of the support vector regression machine can be realized, the parameters in the model have a great relationship. After the selection of the kernel function, we use the enumeration method to select the appropriate kernel function width p=5, as well as the penalty parameter C=500. Then we train the SVR prediction model and get the prediction model of wind speed data.
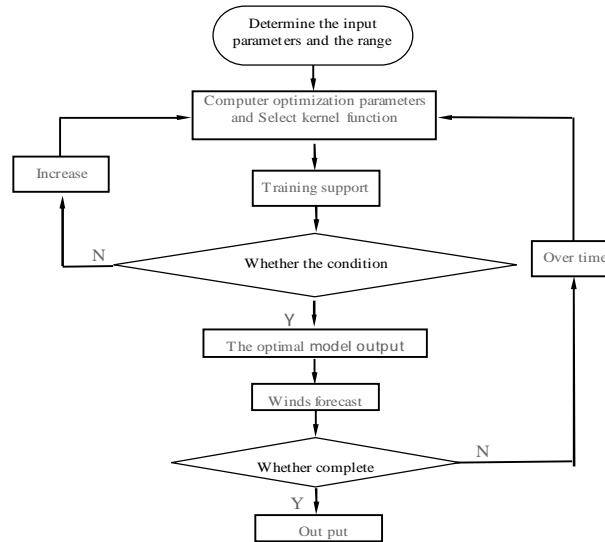
The figure below shows the SVR prediction winds flow chart.

Fig.4 the process of Support vector regression machine forecasting wind speed

We use the trained SVR prediction model to predict the wind speed value of 301-310. The predicted results are shown in Figure 5. In the figure, "O" means the true value, "+" indicates the predicted value.
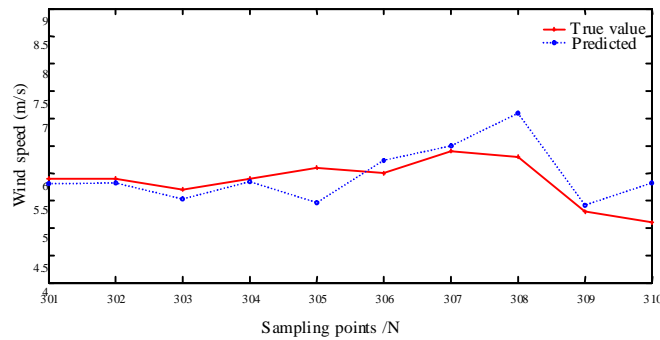

Fig.5   The predictive results of the wind speed data

It can be seen from the figure, the prediction results can well reflect the fluctuation of wind speed, and the difference between predictive value and true value is small. In order to study the prediction results, we use the absolute error, relative error, mean absolute error and mean relative error between the predictive value and the real value to judge the results. The results of the comparison between the predicted value and the real value are shown in Table 1.

As can be seen from the table, the establishment of the prediction model can be very good for the wind speed forecast, and the prediction results are relatively accurate.

Table 1 predicted results contrast with the real data of wind speed

No.	Actual value	Predictive value	Error	Relative error
301	5.9	5.8035	0.0965	0.0164
302	5.9	5.8171	0.0829	0.0141
303	5.7	5.5264	0.1736	0.0305
304	5.9	5.8395	0.0605	0.0103
305	6.1	5.4617	0.6383	0.1046
306	6.0	6.2281	0.2281	0.0380
307	6.4	6.4951	0.0951	0.0149
308	6.3	7.0953	0.7953	0.1262
309	5.3	5.4134	0.1134	0.0214
310	5.1	5.8171	0.7171	0.1406
The sum of absolute error		3.0009	The sum of relative error	0.0517
The average absolute error		0.30009	The average relative error	0.00517

Conclusions

SVR is a kind of machine learning technique based on the structural risk minimization principle, and it has good generalization performance and accuracy. Especially in solving the small sample and nonlinear problems, SVR has its own unique advantages. It can be seen from the data of a certain wind farm, the support vector regression machine can predict the wind speed more accurately. In addition, compared with other existing nonlinear regression methods, it has smaller computation and can solve the problem of regression estimation, such as wind speed forecasting.

Support vector machine regression function is determined only by a small number of support vectors and its complexity depends on the number of support vectors instead of the dimension of sample space, which avoids the "Curse of dimensionality" in a sense and extends the application range of multiple regression.

References

[1] Gu GuoLi,Wang Weiqing, et al. Wind speed forecasting method research [J]. Journal of scientific research.2009,6 (6)：22-24

[2] Liu ye,Lu xiaofang, et al. The wind power system wind speed forecasting methods presented [J]. Power grid and clean energy, 2010, 6 (6) : 62-66

[3] Ding Shifei,Qi BingJuan,Tan hongyan. Support vector machine (SVM) theory and algorithm research review [J]. Journal of electronic science and technology university, 2011,1(1):2-9

[4] Du Shuxin,Wu Tiejun. Used for regression estimates for support vector machine (SVM) method [J]. Journal of system simulation, 2003, 6 (11). 1580-1585

[5] Wang wei,Lin Ming, et al. Support vector machine (SVM) and its application prospect in earthquake prediction [J]. Journal of northwest earthquake, 2006,1 (3)

[6] Wang Qinglan. Statistical learning theory [J]. Journal of tangshan teachers college,2007,5 (9)

[7] S.Amari,S.Wu.Improving support vector machine classifiers by modifying kernel functions[J].IEEE Trans. Neural Networks,1999,12:783~789

[8] CAO Liangyue,Practical method for determining the minimum embedding dimension of a scalar time series [J],Physical D,1997,110 (122):43-50